Discriminating spontaneous locomotor play of dairy calves using accelerometers

Citation for published version:
Größbacher, V, Buková, K, Lawrence, AB, Špinka, M & Winckler, C 2020, 'Discriminating spontaneous locomotor play of dairy calves using accelerometers', *Journal of Dairy Science*, pp. 1866-1873. https://doi.org/10.3168/jds.2019-17005

Digital Object Identifier (DOI):
10.3168/jds.2019-17005

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of Dairy Science

Discriminating spontaneous locomotor play of dairy calves using accelerometers

V. Größbacher[*†1], K. Bučková[‡§], A. B. Lawrence[#], M. Špinka[‡§] and C. Winckler[*]

[*]Department of Sustainable Agricultural Systems, University of Natural Resources and Life Sciences, 1180 Vienna, Austria

[†]Recipient of a DOC Fellowship of the Austrian Academy of Sciences

[‡]Department of Ethology and Companion Animal Science, Czech University of Life Sciences, 165 00 Prague, Czech Republic

[§]Department of Ethology, Institute of Animal Science, 104 00 Praha Uhříněves, Czech Republic

[#]Animal and Veterinary Sciences Research Group, SRUC (Scotland's Rural College), EH9 3JG Edinburgh, United Kingdom

[1]Corresponding author: verena.groessbacher@boku.ac.at

INTERPRETIVE SUMMARY

Title: Discriminating spontaneous locomotor play of dairy calves using accelerometers

First author's last name: Größbacher

In calves, play behavior is a promising indicator to assess both compromised and enhanced welfare. However, quantifying play is difficult due to its rare and irregular occurrence. We aimed to validate 1 Hz accelerometer recordings to measure locomotor play of dairy calves in their home-pens. Accelerometer data were combined into 10 s periods and periods were categorized as PLAY/NOPLAY using quadratic discriminant analysis. Comparing these periods with behavior recorded from video, play was correctly classified in 79% of cases. Based on a correlation of r_P=0.87 with observed play intervals, it may be used as a proxy to replace behavior observations.

ABSTRACT

Play behavior is a promising welfare indicator in dairy calves as it decreases in negative situations such as pain or hunger and increases in positive contexts such as in appropriate social environment. Directly measuring play is time consuming as it is performed in irregular bouts and can be inconsistent over days. To facilitate automatic recording of play, previous studies fitted tri-axial accelerometers to the hind legs of calves, measuring the velocity of movements in large arenas, and reported high correlations between vertical axis peak duration and the duration of locomotor play. The current study aimed at validating accelerometers for recording spontaneous locomotor play in the calves' home-pens over longer periods of time. Data were collected from 48 Holstein Friesian calves at either four or eight weeks of age, housed in groups of three in pens of 10 m^2. Acceleration at the vertical axis of the hind leg was recorded at a rate of 1 Hz. One active time period for each calf was randomly selected (mean duration ± SD = 34 ± 9 min). From video of the corresponding time period, frequency of locomotor play events consisting of run, turn and buck/buck-kick was recorded using behavior sampling. Combined counts of play events were highly correlated (r_p = 0.91) with counts of peaks in acceleration. However, for calves with higher levels of locomotor play, this method underestimated the extent of play. Alternatively, run, turn and buck events obtained from video were transformed into a binary response by creating intervals of 10s and then classifying each 10s interval as comprising events of play (PLAY) or not comprising events of play (NOPLAY). The corresponding accelerometer data for all 10s periods, equaling 10 consecutive readings each, were classified into PLAY or NOPLAY with quadratic discriminant analysis. 79% of periods with locomotor play were correctly classified. Counts of observed play intervals correlated with the counts of play periods from accelerometers at r_p = 0.87, but the discriminant analysis consistently overestimated play. In conclusion, accelerometer measurements at 1 Hz (in 1 s intervals) and at the vertical axis alone cannot be used to exactly quantify absolute levels of locomotor play in the home-pen. However, counts of peak accelerations can provide a rough

estimate of inter-individual differences in play events and discriminant analysis can be used as a proxy for one-zero sampling of inter-individual differences in locomotor play.

Key words: automated measuring, acceleration, behavior classification, dairy calf

INTRODUCTION

In the past decade, accelerometers have found increasing application in farm animal research. The use of accelerometers facilitates data collection as automatic recording can circumvent time and resource intensive behavior observations. In calves, accelerometers have been primarily employed to record general activities. For example lying, standing or locomotion of dairy calves have been recorded using accelerometers to detect early behavioral signs of respiratory diseases (Swartz et al., 2017) and of neonatal diarrhea (Sutherland et al., 2018b). Moreover acceleration measurements have been used to quantify lying and standing when studying effects of social housing on weaning (Overvest et al., 2018) and determining effects of different disbudding methods on lying behavior (Sutherland et al., 2018a). While the accuracy of recording general activities from accelerometers is high, the validation of recording specific behaviors such as feeding and ruminating is still in progress (e.g. Roland et al., 2018).

In calves, play behavior is a promising indicator to assess both compromised welfare, e.g. pain after disbudding (Mintline et al., 2013) or hunger after weaning off milk reduced play (Krachun et al., 2010, Miguel-Pacheco et al., 2015), and enhanced welfare, e.g. group housing increased play (Valníčková et al., 2015). However, calves perform spontaneous play for only a few minutes per day (Jensen et al., 2015) at irregular intervals (Fraser and Duncan, 1998). Thereby quantification of play from observation is usually accomplished either through continuous recording of durations or events (e.g. Jensen et al., 2015, Miguel-Pacheco et al., 2015) or through one-zero sampling of the presence of play in certain sample intervals (e.g. Valníčková et al., 2015). These challenges associated with measuring play behavior raise the interest in automatic recording techniques. In previous studies, accelerometers were used to automatically record locomotor play of calves, however the recordings were conducted in large arenas and for a short time only. Rushen and de Passillé (2012) found correlations of up to $r_s = 0.88$ between the duration of running and the sum of total acceleration in all three axes and Luu et al. (2013) found correlations of up to $r_p = 0.98$ between the duration of locomotor play (running plus jumping/kicking) and the sum of the percent of peaks (3 *g* or higher) of all axes. In both studies acceleration was recorded at a high rate of 33 Hz and in all three axes, limiting the recording duration to 10 min due to the memory capacity of the accelerometers (HOBO Pendant G Acceleration Data Logger, Onset Computer Corporation, Pocasset, MA, USA). With the intention of assessing a longer recording duration, Luu et al. (2013) simulated a sampling frequency of 1 Hz by taking every 33rd acceleration reading on one axis only and found a correlation of $r_p = 0.92$ between the percent of peaks on the vertical axis and the duration of locomotor play. In order to test the practical application of measuring spontaneous locomotor play over the duration of many hours, the current study aimed to validate the use of accelerometers to measure play behavior in the home-pens of calves at low recording rates. Specifically, our objectives were using recordings at 1 Hz on the vertical axis (1) to test whether counts of peak accelerations can accurately estimate events of locomotor play and (2) to determine whether classifying periods of acceleration readings into PLAY/NOPLAY can reliably measure play behavior recorded by one-zero sampling from video.

MATERIAL AND METHODS

All data were collected at the Netluky Research Station at the Institute of Animal Science in Prague/Czech Republic between August 2016 and April 2017. The study was approved by the Institutional Animal Care and Use Committee of the Institute of Animal Science in Prague and the Czech Central Committee for Protection of Animals, Ministry of Agriculture (permit number 27356/2016-MZE-17214).

Animals and Housing

The 48 Holstein Friesian-calves (20 female, 28 male) reported on here were a subset of a larger study using 72 calves. They were housed in an uninsulated barn with wind-shields in 24 groups of three. Pens were 10.1 m^2 with a straw-bedded lying area of 4.2 x 1.4 m and a concrete activity and feeding area of 3.5 x 1.2 m. Calves entered group-housing at an average age of 13.3 ± 3.1 days (mean ± SD) with groups entering the experiment consecutively. Calf allocation to groups was balanced for sex, age and weight. For the purpose of another experiment calves were fed either 6 liters of milk daily throughout the experiment or they received 9 liters per day at week 4 and the provision continuously increased to 12 liters at week 6. All calves received three milk meals per day in teat buckets. All calves received 3 liters of milk in the morning. Calves with a low milk allowance received 1.5 liters of milk per meal at midday and in the evening. Calves with a high milk allowance received 3 liters continually increasing to 4.5 liters of milk per meal at midday and in the evening. Calves had ad libitum access to water, concentrates and hay offered in buckets. Among the 72 calves, two focal calves per group were randomly selected with observations of one calf taking place at 4 weeks and the other at 8 weeks of age. Calves weighed 57.5 ± 5.7 kg (mean ± SD) at 4 weeks and 88.3 ± 12.4 kg at 8 weeks.

Acceleration measurements

Accelerometers (HOBO Pendant G Acceleration Data Logger, Onset Computer Corporation, Pocasset, MA, USA; product specifications are described in detail in Luu et al. (2013)) were fitted to the rear side of both hind legs of calves using elastic cohesive bandages. Accelerometers were attached vertically to the leg such that the x-axis was perpendicular to the ground. The accelerometers were set to measure readings on the vertical axis at a rate of 1 Hz (1 sample/s), allowing recording of acceleration every second for 18.1 hours. Acceleration was recorded from 05.00 until 23.04 on two consecutive days per testing week with the accelerometer on the right leg recording day 1 and on the left leg recording day 2. Calves wore accelerometers for approximately 66 hours per testing week (from the evening before until the morning after the testing days). Programming of accelerometers was performed using an optical infrared base station with USB interface and the HOBOware Pro Software (Version 3.7.8; Onset Computer Corporation, Pocasset, MA, USA) with the starting time set in advance.

Behavior observation

Behavior of calves was video recorded for 48 hours per testing week using one camera per pen (VCC-HD2300P, Sanyo, Japan; FW2220R-Z, Dahua Technology Co., China; HDC-SD99, Panasonic, Japan) and infrared radiators (RM50-AI-50, Raytec, UK; LIR-T80 and LIR-T60, IR LAB Surveillance Tech, Taiwan). Based on the graphic display of downloaded acceleration values, using the plot-function of the HOBOware Pro Software, lying and active phases could be clearly distinguished. Therewith one activity bout of approx. 30 min was selected for each calf. Activity bouts were selected in a time span between 05.00 and 20.00 when accelerometer recordings were available and video recordings allowed easy distinction of behaviors due to daylight hours. The week and day of the selected activity bout was randomized for each focal calf. Selection of activity bouts was balanced across different times of the day and start time of selected bouts ranged from 06.04 until 19.23. The duration of selected activity bouts was 34.3 ± 9.2 min (mean ± SD). For individual recognition calves were marked across their backs and sides with animal marking sticks. Behaviors categorized as locomotor play are described in Table 1. The criterion interval for halts in between running events was set to 1 second based on visual assessment of a log survivorship plot. Events of locomotor play behavior were continuously recorded by one person using the Mangold INTERACT video analysis software (Version 16.1.5.8). Intra-observer-reliability was measured from 3 randomly selected activity bouts of 41.7 ± 8.0 min each (mean ± SD) assessed two times. A Wexler’s ratio was calculated from the number of agreements (i.e. the number of locomotor play events that were correctly scored within one second in both recording sessions x 2) divided by the number of possible agreements (i.e. the total number of locomotor play events scored in both sessions; e.g. used in Wathan et al. 2015). Wexler's ratio was assessed for each activity bout individually, with an average agreement ratio of 0.84. Continuous recording was transformed into one-zero sampling by creating sample intervals of 10 s and classifying them according to presence or absence of locomotor play events within the interval.

Data analysis

All statistical analyses were performed in SAS 9.4. We analyzed the acceleration data according to two methodologies:

Peak acceleration method (PEAK). We used Pearson correlations to assess the strength of association between counts of peak measurements of acceleration and counts of observed locomotor play events. Counts of peak accelerations were calculated for different upper and lower thresholds of acceleration values in steps of 0.1 *g* (e.g. counts of values ≥ 3.2 *g*, 3.1 *g*, 3.0 *g*,… and ≤ -3.2 *g*, -3.1 *g*, -3.0 *g*,…). Pearson correlations of all 1056 combinations (32 thresholds of ≥ 0 *g* times 31 thresholds of < 0 *g*) of counts of peaks applying different upper and lower thresholds with counts of locomotor play were calculated. Therewith, the best combination of threshold values of acceleration to predict locomotor play was identified as counts of peaks of ≥ + 1.6 *g* and ≤ - 3.0 *g*. While the distribution of data was right-skewed and did not visually conform to the assumption of normality for parametric measures of association, the data distribution was unimodal. Three outliers were visually detected using a Cook's Distance plot, though no outlier with leverage was visually identified in the outlier and leverage diagnostics (leverage and studentized residuals).

Classifier method (CLASS). We used quadratic discriminant analysis to predict a categorical response (Kuhlenkasper and Handl, 2017), i.e. the occurrence of locomotor play in each period (10 s fragment of observations) based on classifiers (predictor variables describing acceleration values in each period (James et al., 2015)). As discriminant analysis requires two sets of data, one set to train the discriminant function and one set to test its predictions, we divided the recorded activity bouts in half. Therefore the accelerometer data were combined to 10 s periods, resulting in 10 measurements per period. Subsequently the periods were alternately allocated to a training data set or a testing data set (testing data set: n=48, mean number of periods ± SD = 102.3 ± 26.9). The presence (PLAY) or absence (NOPLAY) of locomotor play in each period was identified from video observation and used as the gold standard. For each period the following metrics were calculated as classifiers derived from the original value (OV) or change in values (CV = $x_i - x_{i-1}$) e.g. minimum, maximum, mean, median, quartiles, variance, total sum; a full list is provided in Supplemental Table S1. Relevant classifiers were then visually preselected from boxplots of PLAY and NOPLAY from the training data set when the interquartile range of NOPLAY was low with little to no overlap with PLAY and when outliers were not widely dispersed. A quadratic discriminant function was then developed with classification probabilities based on the proportional occurrences of how often PLAY and NOPLAY were scored in the training data set, i.e. 97% of periods displaying NOPLAY and 3% of periods displaying PLAY. With the testing data set the predictive abilities of the discriminant function were assessed. Discriminant functions with different combinations of classifiers were tested and the combination of classifiers with the highest sensitivity and specificity was selected. The relevant classifiers included in the final discriminant function are displayed in Table 2. Discriminant analysis assumes a Gaussian distribution from observations of each class (James et al., 2015). The present data of the values of classifiers could not be assumed to be normally distributed nor could data be transformed to fit the underlying assumptions of normality. We were able to circumvent this issue by dividing the data set into two halves, a training data set and a testing data set.

Therewith the performance of the discriminant function was not contingent on the data distribution and could be independently verified.

Comparison of PEAK and CLASS. In order to directly contrast the outcome of the two methodologies on the basis of the same set of data, we calculated PEAK and CLASS with the testing data set only (n = 48; mean duration ± SD = 17.1 ± 4.6 min). To assess the strength of association between the measures of acceleration and the observed locomotor play, a Pearson correlation of counts of peaks resulting from the PEAK method and counts of observed locomotor play events was calculated. Likewise, a Pearson correlation of counts of PLAY periods resulting from the CLASS method with counts of observed locomotor play intervals from one-zero sampling was calculated. In order to assess the magnitude of disagreement and facilitate the detection of trends, we produced Bland-Altman plots. The plots depict the average of the acceleration measure and the observation on the x-axis and the difference between the acceleration measure and the observation on the y-axis (Altman and Bland, 1983). Bland-Altman plots were produced for both methodologies of analysis and compared visually.

RESULTS

When assessing play by continuous recording of frequencies, calves performed 5.3 events of locomotor play per 30 min observation period (SD = 7.3; range = 0 - 27). The Pearson correlation with counts of peaks of ≥ + 3.0 *g* and ≤ - 3.0 *g* from the corresponding accelerometer data, as described by Luu et al. (2013), was 0.83 (P<0.01). However, we attained the highest correlation with counts of locomotor play when using counts of peaks of ≥ + 1.6 *g* and ≤ - 3.0 *g* (r_p = 0.91, P > 0.01; Figure 1). The respective scatter plot (Figure 1) illustrates a strong linear relationship of both measurements, but an unequal rate of increase of counts of peaks with counts of play is noticeable. The Bland-Altman plot (Figure 2) further emphasizes the uneven distribution across the range of locomotor play as higher counts of locomotor play events were increasingly underestimated by the peak acceleration method, demonstrating that the number of play events and the number of accelerometer peaks did not directly correspond to each other, i.e., they are not on the same scale. The mean deviation of peak measurements from observed play events amounted to - 1.90 ± 4.42.

Alternatively, when recording locomotor play with one-zero sampling, calves performed play in 2.7 periods per observation (SD = 3.5; range = 0 - 16). From the accelerometer data, we estimated the number of play periods using the classifier method with the outcome displayed as contingency table (Table 3). It follows that CLASS overestimates the number of PLAY periods. CLASS achieved a precision of 0.95 (= proportion of correctly classified periods), a sensitivity of 0.79 (= proportion of correctly classified true positives) and a specificity of 0.96 (= proportion of correctly classified true negatives). Counts of PLAY periods identified with CLASS highly correlated with counts of observed PLAY periods recorded from video (r_p = 0.87; $P < 0.01$; Figure 3). The scatter plot (Figure 3) illustrates a strong linear relationship of both measurements but indicates an intercept and concomitant overestimation of PLAY periods by the CLASS method. The number of accelerometer-identified PLAY periods surpasses the number recorded visually by 3.65 ± 2.42 periods; nonetheless the Bland-Altman plot (Figure 4) shows an evenly distributed deviation of the two measurements across the range of counts of PLAY periods.

DISCUSSION

With this study we aimed at providing an approach to automatically record locomotor play of calves in their home-pen and for long durations using acceleration measurements. In previous studies accelerometers have been validly used to record durations of lying and standing in calves (Bonk et al., 2013, Swartz et al., 2016). Similarly in the current study we were able to easily distinguish between lying and standing on the vertical axis, with values of lying fluctuating around 0 g and values of standing around - 1 g, depending on the position of the hind leg. Therefore with - 1 g as the center of fluctuation, measuring play with peaks of $\geq$ + 1.6 g and $\leq$ - 3.0 g is sensible. We reason that peaks had not reached + 3.0 g, as reported by Luu et al. (2013), because the smaller dimensions of the home-pens in comparison with a large arena did not permit calves to consistently reach accelerations of a similarly high level. Thus small spaces can restrict the magnitude of movement and also fragment the occurrence of play (Jensen et al. 1998). Nevertheless we cannot draw conclusions on any space allowance between our home-pen and the arena of Luu et al. (2013) as this was not part of our investigation. Moreover locomotor play consists of rapid motions of the hind legs for short durations and is often nested within short time intervals. Therefore recordings at 1 Hz and on one axis may be too infrequent to accurately capture locomotor play events in the home-pen, resulting in the unequal increase and accretive underestimation of higher frequencies of play events of the PEAK-method, as visualized in the Scatter plot and Bland-Altman plot. Nevertheless, the high correlation of peak accelerations and observed play events elucidates a strong link between the two recording methods. Thus, while the PEAK-method cannot record the duration of play in absolute terms, it can produce an approximate estimation of play levels and allows the comparison of relative differences between calves in standard housing conditions.

In the CLASS method, we used the accelerometer data to simulate the one-zero observational method by merging the recordings to 10 s periods, thus ensuring the use of repeated measures and circumventing the need to count individual peaks above/below a certain threshold. This allowed us to view acceleration values in context, integrated with the values preceding and following them. We derived classifiers from combined values e.g. mean of two highest values or variance to mathematically describe the 10 acceleration measures per period and highlight the differences between PLAY and NOPLAY. Thereby we classified brief time spans according to the presence or absence of locomotor play within these 10 seconds. The use of original individual acceleration values e.g. the mere minimum or maximum value would have resulted in a lower sensitivity to correctly identify PLAY periods. Such an approach has been previously successfully implemented in accelerometer validation regarding sheep gait, describing periods with relative frequencies of integers e.g. the number of high frequency acceleration readings between - 4 and - 3 per period divided by all readings of the period (Radeski and Ilieski, 2017). Other studies described periods using movement metrics e.g. mean, variance and inverse coefficient of variation (Watanabe et al., 2008) or signal magnitude area, average intensity and average entropy (Barwick et al., 2018). However, in these studies acceleration was recorded at a higher rate. Recording at a higher rate would have also allowed for classifying shorter periods. For example Radeski and Ilieski (2017) recorded at 33 Hz and classified periods of three seconds.

CLASS correctly discriminated 79% of periods with an occurrence of locomotor play, but at the same time overestimated play by approximately 200% (out of 304 play periods identified by CLASS, 102 periods were true positives and 202 periods were false positives). Similarly other accelerometer models consistently overestimated locomotor behavior e.g. Swartz et al. (2016) overestimated stepping by 18% and Trénel et al. (2009) consistently overestimated moving activity with a ratio of probability of correct negatives to correct positives of 7.57 (PV- = 0.98, PV+ = 0.13). Thus while the number of classified play periods is strongly associated with counts of observed play intervals, overall the classifier method overestimates locomotor play in absolute terms and produces an intercept by adding 3.7 play periods to each observation. However, the Bland-Altmann plot shows a rather consistent and evenly distributed deviation across the range of number of play intervals observed without indication of a directed effect. Hence, while CLASS cannot accurately measure locomotor play, it can be used as a proxy. After factoring in the consistent overestimation, it assesses the number of play periods close to the scale of one-zero sampling and thus allows comparing absolute differences between individuals. With these results we offer a feasible approach to assess spontaneous locomotor play in home-pens of calves using an affordable and commercially available accelerometer model for durations of many hours or perhaps even days. Nevertheless these results may only be valid for the housing conditions investigated and further studies are needed to validate this approach under e.g. different space allowances. A prerequisite to classify periods with discriminant analysis is to use shorter subsets of behavior recordings as a training data set. In the current study only active periods of animals were included in the analysis. In order to apply the classifier method to the full data set it is necessary to either preselect only active periods or to include lying bouts in the training data set. Therewith it is feasible to train the discriminant function with the selected classifiers and thereafter to apply it to the entire recordings of acceleration.

We must stress that our proposed approach with recording at a frequency of 1 Hz can only be used as an approximate estimation of locomotor play. A higher level of accuracy could be achieved by increasing the rate of recording. Measuring acceleration at the highest rate (33 Hz) allowed de Passillé et al. (2010) to measure the interstep interval and accurately distinguish between different gait patterns. Radeski and Ilieski (2017) were able to achieve high accuracy in classifying 3 s periods of walking, trotting and galloping in sheep with discriminant analysis, when recorded at a rate of 33 Hz. In the current study the recording rate was limited by its data storage capacity, however Le Roux et al. (2018) achieved a 469-fold reduction in memory requirement when classifying lying, standing and walking on the accelerometer rather than storing raw data. Thus, the proposed approach is easily applicable and inexpensive with the available resources, however there are numerous options to improve the accuracy of recording by availing technical advancements.

CONCLUSION

Using the peak acceleration method, the acceleration of calves' hind legs measured at a rate of 1 Hz can be used to obtain an approximate estimation of inter-individual differences in the occurrence of locomotor play events. Quadratic discriminant analysis can replace observational one-zero sampling, when based on indirect movement metrics obtained from 10-second-periods of raw accelerometer data. This alternative method may be more accurate in quantifying the inter-individual differences in locomotor play of dairy calves in their home-pens as it reveals less biased estimates across different levels of play. If the accurate measurement of absolute levels of behavior is the ultimate aim of automatic recording, a sensor with higher memory capacity must be found.

ACKNOWLEDGEMENTS

We are grateful to the farm staff at the Netluky Research Station for the care of calves and cooperation with the research team. This study was funded by the Austrian Academy of Sciences and by the grant MZE-RO0718 from the Czech Ministry of Agriculture.

REFERENCES

Altman, D. G. and J. M. Bland. 1983. Measurement in Medicine: the Analysis of Method Comparison Studies. Statistician 32:307-317.

Barwick, J., D. W. Lamb, R. Dobos, M. Welch, and M. Trotter. 2018. Categorising sheep activity using a tri-axial accelerometer. Comput Electron Agric 145:289-297. https://doi.org/10.1016/j.compag.2018.01.007

Bonk, S., O. Burfeind, V. S. Suthar, and W. Heuwieser. 2013. Technical note: Evaluation of data loggers for measuring lying behavior in dairy calves. J. Dairy Sci. 96:3265-3271. https://doi.org/10.3168/jds.2012-6003

de Passillé, A. M., M. B. Jensen, N. Chapinal, and J. Rushen. 2010. Technical note: Use of accelerometers to describe gait patterns in dairy calves. J. Dairy Sci. 93:3287-3293. https://doi.org/10.3168/jds.2009-2758

Fraser, D. and I. J. H. Duncan. 1998. 'Pleasures', 'pains' and animal welfare: Toward a natural history of affect. Anim. Welfare 7:383-396.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2015. An Introduction to Statistical Learning with Applications in R. Springer Science+Business Media, New York. https://doi.org/10.1007/978-1-4614-7138-7

Jensen, M. B., L. R. Duve, and D. M. Weary. 2015. Pair housing and enhanced milk allowance increase play behavior and improve performance in dairy calves. J. Dairy Sci. 98:2568-2575. https://doi.org/10.3168/jds.2014-8272

Jensen, M. B. and R. Kyhn. 2000. Play behaviour in group-housed dairy calves, the effect of space allowance. Appl. Anim. Behav. Sci. 67:35-46. https://doi.org/10.1016/S0168-1591(99)00113-6

Jensen, M. B., K. S. Vestergaard, and C. C. Krohn. 1998. Play behaviour in dairy calves kept in pens: the effect of social contact and space allowance. Appl. Anim. Behav. Sci. 56:97-108. https://doi.org/10.1016/S0168-1591(97)00106-8

Krachun, C., J. Rushen, and A. M. de Passillé. 2010. Play behaviour in dairy calves is reduced by weaning and by a low energy intake. Appl. Anim. Behav. Sci. 122:71-76. https://doi.org/10.1016/j.applanim.2009.12.002

Kuhlenkasper, T. and A. Handl. 2017. Diskriminanzanalyse. Pages 363-413 in Multivariate Analysemethoden. Theorie und Praxis mit R. 3rd ed. Springer Spektrum, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-54754-0_12

Le Roux, S. P., R. Wolhuter, N. Stevens, and T. Niesler. 2018. Reduced Energy and Memory Requirements by On-Board Behavior Classification for Animal-Borne Sensor Applications. IEEE Sens J 18:4261-4268. https://doi.org/10.1109/jsen.2018.2816965

Luu, J., J. F. Johnsen, A. M. d. Passillé, and J. Rushen. 2013. Which measures of acceleration best estimate the duration of locomotor play by dairy calves? Appl. Anim. Behav. Sci. 148:21-27. https://doi.org/10.1016/j.applanim.2013.07.004

Miguel-Pacheco, G. G., A. Vaughan, A. M. de Passille, and J. Rushen. 2015. Relationship between locomotor play of dairy calves and their weight gains and energy intakes around weaning. Animal 9:1038-1044. https://doi.org/10.1017/S1751731115000063

Mintline, E. M., M. Stewart, A. R. Rogers, N. R. Cox, G. A. Verkerk, J. M. Stookey, J. R. Webster, and C. B. Tucker. 2013. Play behavior as an indicator of animal welfare: Disbudding in dairy calves. Appl. Anim. Behav. Sci. 144:22-30. https://doi.org/10.1016/j.applanim.2012.12.008

Overvest, M. A., R. E. Crossley, E. K. Miller-Cushon, and T. J. DeVries. 2018. Social housing influences the behavior and feed intake of dairy calves during weaning. J. Dairy Sci. 101:8123-8134. https://doi.org/10.3168/jds.2018-14465

Radeski, M. and V. Ilieski. 2017. Gait and posture discrimination in sheep using a tri-axial accelerometer. Animal 11:1249-1257. https://doi.org/10.1017/s175173111600255x

Roland, L., V. Schweinzer, P. Kanz, G. Sattlecker, F. Kickinger, L. Lidauer, A. Berger, W. Auer, J. Mayer, V. Sturm, D. Efrosinin, S. Breitenberger, M. Drillich, and M. Iwersen. 2018. Technical note: Evaluation of a triaxial accelerometer for monitoring selected behaviors in dairy calves. J. Dairy Sci. 101:10421-10427. https://doi.org/10.3168/jds.2018-14720

Rushen, J. and A. M. de Passillé. 2012. Automated measurement of acceleration can detect effects of age, dehorning and weaning on locomotor play of calves. Appl. Anim. Behav. Sci. 139:169-174. https://doi.org/10.1016/j.applanim.2012.04.011

SAS Institute. 2012. User's Guide: Statistics, Version 9.4 Edition. SAS Institute, Inc., Cary, NC.

Sutherland, M. A., J. Larive, V. Cave, and G. Zobel. 2018a. Behavioural and physiological responses to clove oil injected under the horn bud of calves. Appl. Anim. Behav. Sci. 204:29-36. https://doi.org/10.1016/j.applanim.2018.03.004

Sutherland, M. A., G. L. Lowe, F. J. Huddart, J. R. Waas, and M. Stewart. 2018b. Measurement of dairy calf behavior prior to onset of clinical disease and in response to disbudding using automated calf feeders and accelerometers. J. Dairy Sci. 101:8208-8216. https://doi.org/10.3168/jds.2017-14207

Swartz, T. H., A. N. Findlay, and C. S. Petersson-Wolfe. 2017. Short communication: Automated detection of behavioral changes from respiratory disease in pre-weaned calves. J. Dairy Sci. 100:9273-9278. https://doi.org/10.3168/jds.2016-12280

Swartz, T. H., M. L. McGilliard, and C. S. Petersson-Wolfe. 2016. Technical note: The use of an accelerometer for measuring step activity and lying behaviors in dairy calves. J. Dairy Sci. 99:9109-9113. https://doi.org/10.3168/jds.2016-11297

Trénel, P., M. B. Jensen, E. L. Decker, and F. Skjøth. 2009. Technical note: Quantifying and characterizing behavior in dairy calves using the IceTag automatic recording device. J. Dairy Sci. 92:3397-3401. https://doi.org/10.3168/jds.2009-2040

Valníčková, B., I. Stěhulová, R. Šárová, and M. Špinka. 2015. The effect of age at separation from the dam and presence of social companions on play behavior and weight gain in dairy calves. J. Dairy Sci. 98:5545-5556. https://doi.org/10.3168/jds.2014-9109

Watanabe, N., S. Sakanoue, K. Kawamura, and T. Kozakai. 2008. Development of an automatic classification system for eating, ruminating and resting behavior of cattle using an accelerometer. Grassl. Sci. 54:231-237. https://doi.org/10.1111/j.1744-697X.2008.00126.x

Wathan, J., A. M. Burrows, B. M. Waller, and K. McComb. 2015. EquiFACS: The Equine Facial Action Coding System. PLoS ONE 10:e01317318. https://doi.org/10.1371/journal.pone.0137818

Table 1. Ethogram of locomotor play derived from Jensen et al. (1998) and Jensen and Kyhn (2000)

Type of locomotor play	Description
Running	Rapid gait with phase of suspension in the air. Minimum of 2 consecutive suspension movements in a forward direction. Running is counted as a new event after 1 second break.
Turning	The two forelegs are lifted from the ground and stretched forward, as the forepart of the body is elevated and turned to one side. Movement upwards and sideward for a minimum of 90 degrees. Occurrence is scored during running bouts.
Bucking/Buck-kicking	Simultaneous lifting of hind legs, claws are raised to a level as high as, or higher than tarsal joints in a standing position. One or both hind legs may be kicked in a posterior or lateral direction. Occurrence is scored during running bouts.

Table 2. Descriptions and equations of the classifiers included in the final discriminant function. Means ± standard deviation of the classifiers are shown for the periods of the testing data set identified as PLAY or NOPLAY from video. OV = original values, CV = change in values, PLAY = period with presence of locomotor play, NOPLAY = period with absence of locomotor play

Classifier	Equation	NOPLAY	PLAY
OV: Mean of two highest acceleration measurements	$\frac{\max(x) + \max2(x)}{2}$	-0.89 ± 0.21	0.04 ± 0.99
OV: Mean of two lowest acceleration measurements	$\frac{\min(\mathrm{x}) + \min2(x)}{2}$	-1.06 ± 0.20	-1.85 ± 0.70
OV: Variance	$\frac{\sum(x-\mu)^2}{10}$	0.03 ± 0.10	0.74 ± 0.80
CV: Maximum of absolute value of change in acceleration measurements	$\max(\lvert\Delta x\rvert)$	0.26 ± 0.48	2.41 ± 1.45
CV: Mean of change in acceleration measurements	$\frac{1}{10}\sum \Delta x_i$	-0.00 ± 0.03	-0.01 ± 0.12
CV: Total sum of absolute values of change in acceleration measurements	$\sum \lvert\Delta x_i\rvert$	0.66 ± 1.15	6.96 ± 4.78

Table 3. Contingency table with number of periods identified with the classifier method (CLASS) as PLAY (event of locomotor play occurring in this period) and NOPLAY (no event of locomotor play occurring in this period)

Observed behavior (Video)	**Predicted behavior (CLASS)**		
	NO PLAY	**PLAY**	**Sum**
NOPLAY	4591	202	4793
PLAY	27	102	129
Sum	4618	304	4922

Figure 1. Relationship between counts of peak accelerations (≥ 1.6 *g* and ≤ 3.0 *g*; PEAK) and counts of locomotor play events observed from video (n = 48 calves). Jitter function was used in the graph to make multiple identical values more visible

Größbacher Figure 1

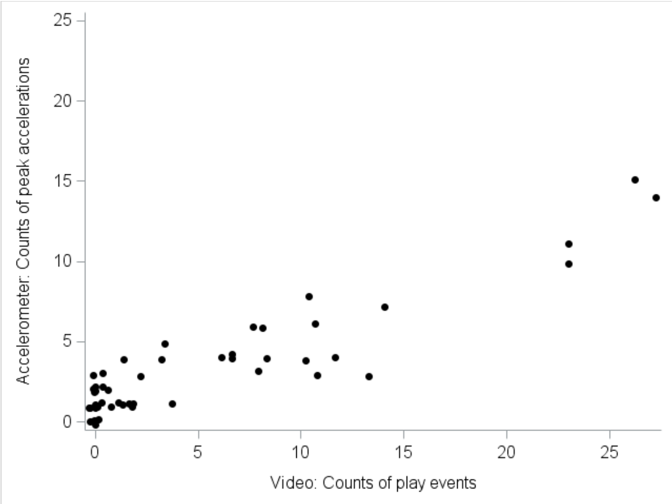


Figure 2. Bland-Altman Plot of the difference in the assessment of locomotor play recorded with accelerometers and video observation compared with the mean of both assessments (PEAK = Peak acceleration method; n = 48 calves). Confidence intervals were estimated at 6.8 at the 95% upper limit and - 10.6 at the 95% lower limit. Jitter function was used in the graph to make multiple identical values more visible

Größbacher Figure 2

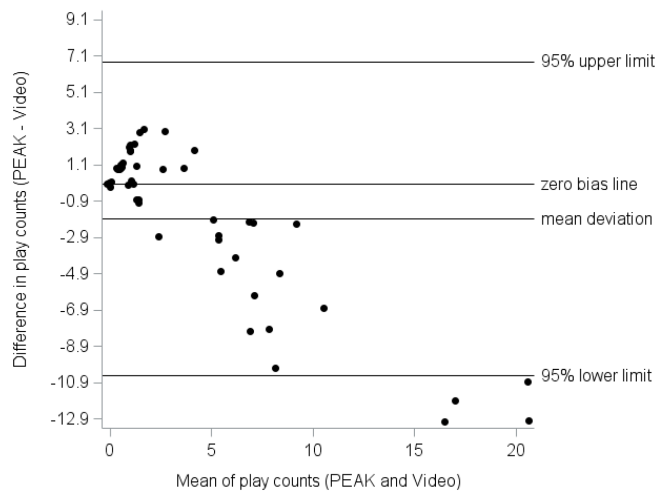


Figure 3. Relationship between counts of periods with locomotor play identified with accelerometers (CLASS = classifier method) and counts of sample intervals with locomotor play observed from video (n = 48 calves). Jitter function was used in the graph to make multiple identical values more visible

Größbacher Figure 3

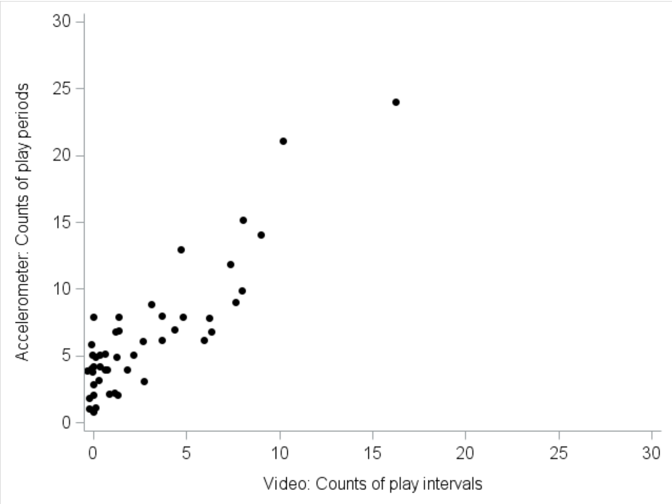
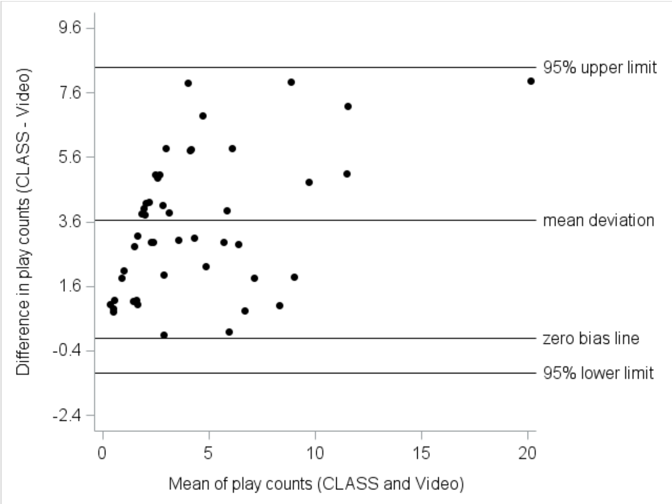


Figure 4. Bland-Altman Plot of the difference in the assessment of locomotor play periods identified with accelerometers and locomotor play intervals from video observation compared with the mean of both assessments (CLASS = Classifier method; n = 48 calves). Confidence intervals were estimated at 8.4 at the 95% upper limit and - 1.1 at the 95% lower limit. Jitter function was used in the graph to make multiple identical values more visible

Größbacher Figure 4



Supplemental Table S1. Descriptions and equations of all potential classifiers not included in the final discriminant function. Means ± standard deviation of the classifiers are shown for the periods of the testing data set identified as PLAY or NOPLAY from video. OV = original values, CV = change in values, PLAY = period with presence of locomotor play, NOPLAY = period with absence of locomotor play. Potential classifiers that were preselected and tested but not included in the final discriminant function are marked with 'Yes'

Classifier	Equation	NOPLAY	PLAY	Preselection
OV: Highest acceleration measurement	$\max(x)$	-0.84 ± 0.33	0.44 ± 1.30	Yes
OV: Second highest acceleration measurement	$\max2(x)$	-0.94 ± 0.12	-0.37 ± 0.84	Yes
OV: Mean of two highest acceleration measurements	$\frac{\max(x) + \max2(x)}{2}$	-0.89 ± 0.21	0.04 ± 0.99	Yes
OV: Third quartile of acceleration measurements	$x_{Q0.75}$	-0.96 ± 0.08	-0.64 ± 0.50	No
OV: Mean of acceleration measurements	$\frac{1}{10}\sum x_i$	-0.98 ± 0.07	-0.93 ± 0.29	No
OV: First quartile of acceleration measurements	$x_{Q0.25}$	-1.00 ± 0.07	-1.24 ± 0.41	No
OV: Mean of two lowest acceleration measurements	$\frac{\min(\mathrm{x}) + \min2(x)}{2}$	-1.06 ± 0.20	-1.85 ± 0.70	Yes
OV: Second lowest acceleration measurement	$\min2(x)$	-1.02 ± 0.10	-1.52 ± 0.69	Yes
OV: Lowest acceleration measurement	$\min(x)$	-1.11 ± 0.33	-2.19 ± 0.86	Yes
OV: Variance	$\frac{\sum(x-\mu)^2}{10}$	0.03 ± 0.10	0.74 ± 0.80	Yes
OV: Total sum of absolute values of acceleration measurements	$\sum\lvert x_i\rvert$	9.84 ± 0.63	11.32 ± 2.20	Yes

CV: Highest change in acceleration measurements	$\max(\Delta x)$	0.22 ± 0.43	2.01 ± 1.34	No
CV: Second highest change in acceleration measurements	$\max 2(\Delta x)$	0.07 ± 0.14	0.86 ± 0.81	No
CV: Third quartile of change in acceleration measurements	$\Delta x_{Q0.75}$	0.04 ± 0.07	0.50 ± 0.55	No
CV: Mean of change in acceleration measurements	$\frac{1}{10}\sum \Delta x_i$	-0.00 ± 0.03	-0.01 ± 0.12	Yes
CV: First quartile of change in acceleration measurements	$\Delta x_{Q0.25}$	-0.04 ± 0.08	-0.49 ± 0.53	No
CV: Second lowest change in acceleration measurement	$\min 2(\Delta x)$	-0.07 ± 0.15	-0.88 ± 0.82	Yes
CV: Minimum of change in acceleration measurement	$\min(\Delta x)$	-0.23 ± 0.44	-2.07 ± 1.33	Yes
CV: Variance of change in acceleration measurements	$\frac{\sum(\Delta x - \mu)^2}{10}$	0.05 ± 0.21	1.56 ± 1.75	Yes
CV: Maximum of absolute value of change in acceleration measurements	$\max(\lvert\Delta x\rvert)$	0.26 ± 0.48	2.41 ± 1.45	Yes
CV: Total sum of absolute values of change in acceleration measurements	$\sum \lvert\Delta x_i\rvert$	0.66 ± 1.15	6.96 ± 4.78	Yes